Exhibit I

EXHIBIT 38

EXHIBIT FILED UNDER SEAL

UBER

Sexual Assaults: Trends + Correlates

Safety & Insurance

Sunny Jeon | Data Scientist
sjeon@uber.com

July 12, 2017

Overview

To support ongoing efforts to prevent sexual misconduct on the Uber platform, this document describes high-level trends + correlates related to the timing, location, and perpetrators of sexual assaults in the US. The results are intended to update previous findings from Feb 2017 (deck available here) using more recent data.

The new results are consistent with previous findings. Sexual assaults are more common late night, on weekends, in spring months, and near bars. They are also more likely to be caused by males, those with a previous history of safety incidents, and those with higher 1-star rates.

These results are described in more detail in the following sections:

1. Definitions [back to top]

The focus of this report is on **sexual assaults**, which are defined here as any kind of non-consensual touching and non-consensual intercourse. These incidents are identified using support tickets that are escalated to incident response teams, which are housed in JIRA. Table 1 describes the different types of incidents within the broader 'sexual misconduct' JIRA category and the percent classified as 'sexual assault' between the July 1, 2016 - July 1, 2017 time period.

Query available here: [redacted]

Table 1: Sexual Assault and Sexual Misconduct Counts in the US (July 2016 - July 2017)

Type	Specific Incident Type	Freq	Perc
Sexual Assault	**Non-Consensual Touching**	**15,681**	**64.29%**
Misconduct	Sexually Inappropriate Remarks or Conversation	4,643	19.03%
Misconduct	Indecent Exposure	1,569	6.43%
Misconduct	Masturbation	824	3.38%
Misconduct	Explicit Gesture	705	2.89%
Misconduct	Staring or Leering	540	2.21%
Sexual Assault	**Non-Consensual Intercourse**	**430**	**1.76%**
		24,392	100%

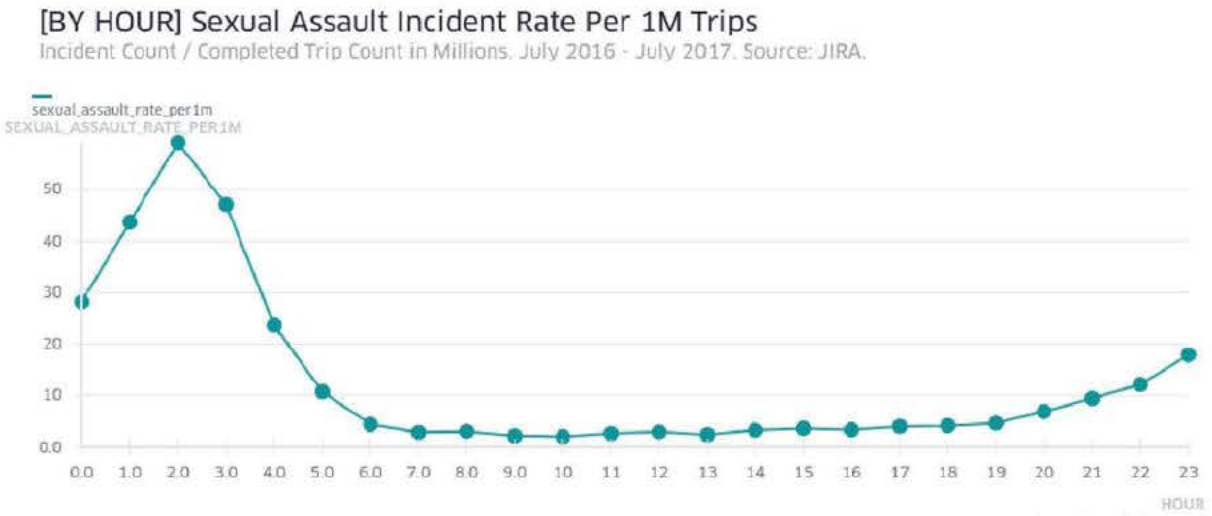
Source: JIRA.

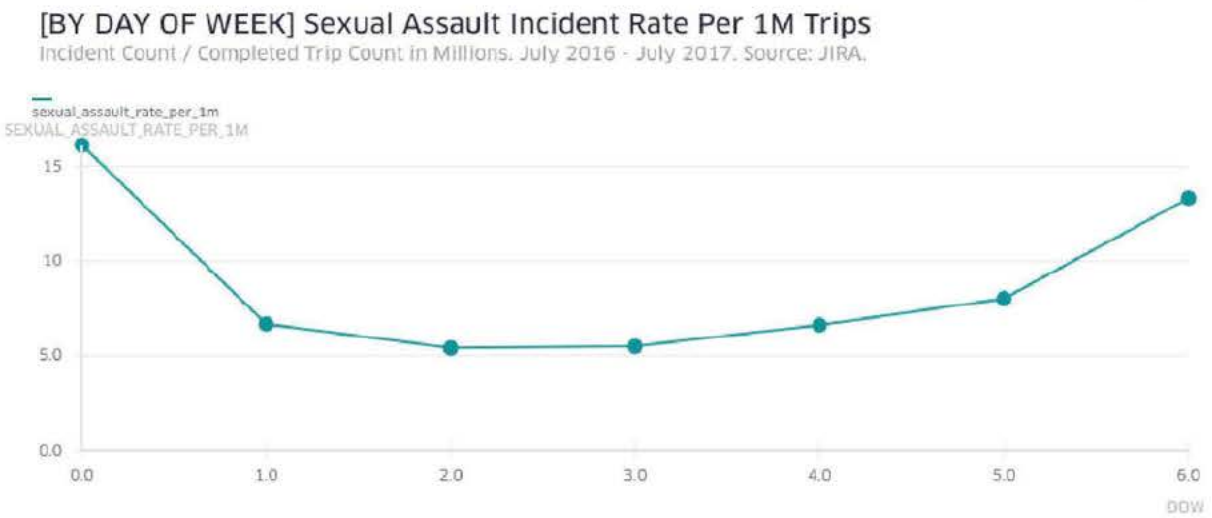
2. High Risk Times, Days, Months [back to top]

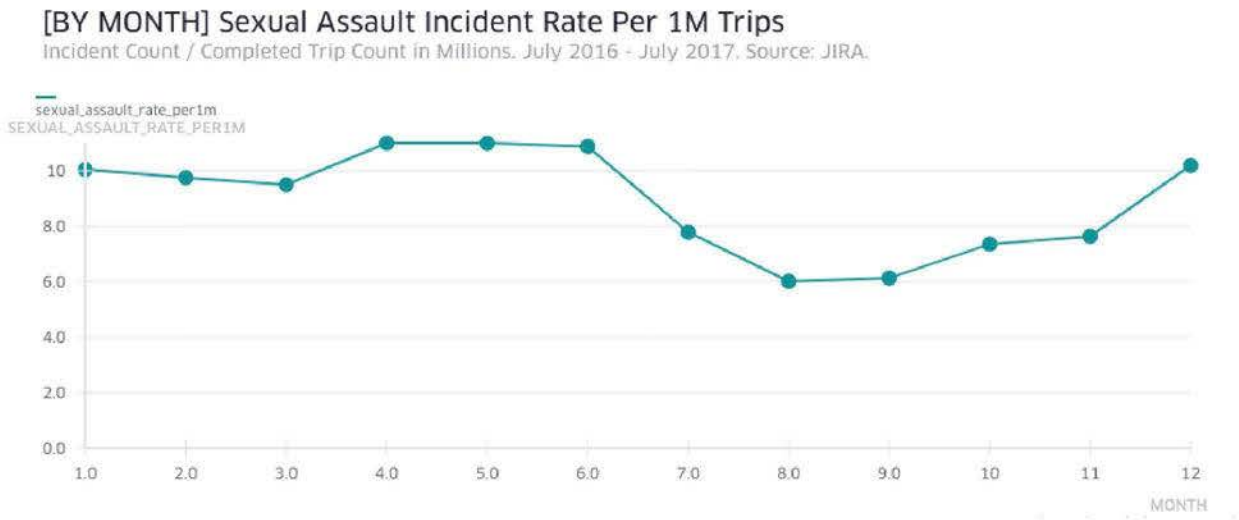
Summary: In the US, the rate of sexual assaults increases late night, on weekends, and in Spring months.

Data / Query:

- Sexual assaults identified using JIRA.
- Data from US cities during the July 1, 2016 - July 1, 2017 time period.
- Queries:
  - Rate by hour: [redacted]
  - Rate by day of week: [redacted]
  - Rate by month: [redacted]







3. Effect of Gender [back to top]

Summary: Of all sexual assaults occurring in the US between Jan, 1 2017 - July 11, 2017, 84% of them had a male offender. When the offender is male, the offender is a driver about half of the time. When the offender is female, however, the offender is more often a rider. Trips that ended in sexual assaults are 2.27x more likely to have a driver and rider of different genders.

Data / Query:

- Sexual assaults identified using JIRA.
- Gender of drivers and rider classified using this table.
- Case-control sampling: all cases (positives) in US from Jan, 1 2017 - July 11, 2017, 1M randomly sampled trips in US during same time period.

Table 2: Gender of Sexual Assault Offenders (US Incidents, Jan - July 2017)

Male	Female	NA
84% (4583 Incidents)	15% (804 Incidents)	1% (76 Incidents)

Table 3: Gender + Role of Sexual Assault Offenders (US Incidents, Jan - July 2017)

Male Offender		Female Offender	
Driver	Rider	Driver	Rider
41.70% (2278 Incidents)	42.19% (2305 Incidents)	1.70% (93 Incidents)	13.01% (711 Incidents)

Table 4: Effects of Different Driver-Rider Genders (Case-Control Sample)

Different Driver-Rider Gender	Cases (Sexual Assaults)	Controls (No Sexual Assault)
Exposed (= 1)	3,919	447,121
Not Exposed (= 0)	2,287	592,875

Percent different gender in cases (sexual assaults): 63%
Percent different gender in controls (no sexual assaults): 43%

Odds Ratio: (3,919/2,287) / (447,121/592,875) = 2.27
95% Confidence Interval: 2.16 - 2.39.

See Appendix A for details on computing and interpreting odds ratios.

4. Proximity to Bars [back to top]

Summary: US trips that end in sexual assaults are 3x more likely to be requested from geohashes where the number of bars is greater than the P99 across all request geohashes.

Data / Query:

- Sexual assaults identified using JIRA.
- Locations of bars identified using the map_services.places table.
- Case-control sampling: all cases (positives) in US from Jan, 1 2017 - July 11, 2017, 1M randomly sampled trips in US during same time period.

Table 5: Number of Bars in Request Geohash (US Incidents, Jan - July 2017)

Geohash 7					
	P25	P50	Mean	P75	P99
Sexual Assaults	0	0	0.59	1	7
No Sexual Assault	0	0	0.33	0	5
Geohash 8					
	P25	P50	Mean	P75	P99
Sexual Assaults	0	0	0.06	0	1
No Sexual Assault	0	0	0.03	0	1

Table 6: Odds Ratios for Number of Bars in Request Geohash

Spatial Unit	Risk Factor	Odds Ratio	95% Confidence Interval
Geohash 7	Num. Bars > Mean	1.772	1.676 - 1.874

Geohash 7	Num. Bars > P99	2.976	2.484 - 3.566
Geohash 8	Num. Bars > Mean	2.365	2.124 - 2.634
Geohash 8	Num. Bars > P99	3.003	2.177 - 4.143

See Appendix A for details on computing and interpreting odds ratios.

5. Previous Safety Incidents [back to top]

Summary: Sexual assaults caused by drivers are 12x more likely to be caused by a driver with >0 previous sexual misconduct complaints against them, and sexual assaults caused by riders are 3x more likely to be caused by a rider with >0 previous sexual misconduct complaints against them. In addition, offenders are 2-3x more likely to have >0 previous IPC tickets.

Data / Query:

- Sexual assaults and previous safety incidents identified using JIRA.
- Case-control sampling: all cases (positives) in US from Jan, 1 2017 - July 11, 2017, 1M randomly sampled trips in US during same time period.

Table 7: Percent with Previous Safety Incidents in Cases and Controls

Sexual Assaults with Driver Offenders		
	Cases (Sexual Assaults)	Controls (No Sexual Assaults)
Percent Drivers with >0 Previous IPC Incident	49.16%	34.55%
Percent Drivers with >0 Previous SA Incident	23.53%	8.41%
Sexual Assaults with Rider Offenders		
	Cases (Sexual Assaults)	Controls (No Sexual Assaults)
Percent Riders with >0 Previous IPC Incident	7.26%	2.3%
Percent Riders with >0 Previous SA Incident	1.82%	0.15%

Table 8: Odds Ratios for Previous Safety Incidents

Offender	Risk Factor	Odds Ratio	95% Confidence Interval
Driver	>0 Previous IPC Incidents	1.832	1.69 - 1.986
Driver	>0 Previous SA Incidents	3.350	3.047 - 3.684
Rider	>0 Previous IPC Incidents	3.325	2.9 - 3.812
Rider	>0 Previous SA Incidents	11.939	9.125 - 15.621

See Appendix A for details on computing and interpreting odds ratios.

6. Feedback from Riders [back to top]

Summary: Comments about drivers from riders contain signals for predicting sexual assaults. Trips ending in driver-caused sexual assaults are nearly 15x more likely to have a driver who previously received rider feedback containing the word "kiss", 5x more likely to have a driver who previously received rider feedback containing the word "sex", and 4x more likely to have a driver who previously received rider feedback containing the phrase "inappropriate behavior".

Data / Query:

- Sexual assaults identified using JIRA.
- Case-control sampling: all cases (positives) in US from Jan, 1 2017 - July 11, 2017, 1M randomly sampled trips in US during same time period.

Table 9: Odds Ratio for "Creepy" Driver Feedback

	Cases	Controls		
	Percent Exposed	Percent Exposed	Odds Ratio	95% CI
Kiss	0.17%	0.01%	14.836	5.473 - 40.219
Sex	0.67%	0.15%	4.526	2.761 - 7.419
Inappropriate Behavior	1.18%	0.33%	3.614	2.486 - 5.254
Flirt	0.5%	0.14%	3.561	2.015 - 6.294
Creepy	2.1%	0.96%	2.207	1.667 - 2.923
Rude	4.16%	2.96%	1.424	1.164 - 1.742

See Appendix A for details on computing and interpreting odds ratios.

7. 1-Star Ratings [back to top]

Summary: Sexual assaults caused by drivers are nearly 3x more likely to be caused by a driver with a 1-star rate greater than the mean, and sexual assaults caused by riders are 1.5x more likely to be caused by a rider with a 1-star rate greater than the mean.

Data / Query:

- Sexual assaults and previous safety incidents identified using JIRA.
- Case-control sampling: all cases (positives) in US from Jan, 1 2017 - July 11, 2017, 1M randomly sampled trips in US during same time period.

Table 10: Odds Ratio for 1-Star Ratings

	Cases	Controls		
	Percent Exposed	Percent Exposed	Odds Ratio	95% CI
Driver Offenders				
1 Star Ratings Prop > Mean	61.01%	36.29%	2.747	2.530 - 2.983
Rider Offenders				
1 Star Ratings Prop > Mean	39.48%	31.10%	1.445	1.344 - 1.553

See Appendix A for details on computing and interpreting odds ratios.

Appendix A: Odds Ratios [back to top]

As a starting point for understanding the data and how various predictors correlate with sexual assaults, I examine *odds ratios* for each predictor, which give the *relative risk* of a sexual assault from exposure to each predictor. The use of a case-control study design in combination with odds ratios here is a common approach used in epidemiology to study risk factors associated with rare diseases.

Table A.1: Cross-Tabulated Data

Predictor	Case	Control
Exposed (= 1)	a	b
Not Exposed (= 0)	c	d

This method involves building a series of 2x2 tables for the case-control data. Table A.1 provides an example. The columns denote cases and controls, and rows represent exposure to a binary predictor, such as whether a trip is a cash trip or not, whether a trip is after midnight, or whether there were >10 rider cancellations in the past 5 minutes, for example. The ratio a/c thus gives the odds of exposure among cases, and b/d gives the odds of exposure among controls. The ratio of these two ratios gives the odds ratio (OR), which represents the relative risk of the outcome (rider-caused incidents) associated with exposure. Substantively, it tell us how much more/less likely the cases (rider-caused incidents) are to have been exposed to a particular risk factor (predictor) compared to controls. More formally, the odds ratio is:

$$OR = (a/c) / (b/d).$$

The confidence interval for the OR parameter is given by:

$$e\text{^}\{\ln OR \pm z * SE_{\ln OR}\},$$

where e is the base on the natural logarithms, z is the confidence coefficient (from the standard normal distribution: $z = 1.96$ for 95% confidence, $z = 2.576$ for 99% confidence), and $SE_{\ln OR}$ is the standard error, which is given by the equation:

$$SE_{\ln OR} = (1/a + 1/b + 1/c + 1/d)^{\wedge 1/2}.$$